

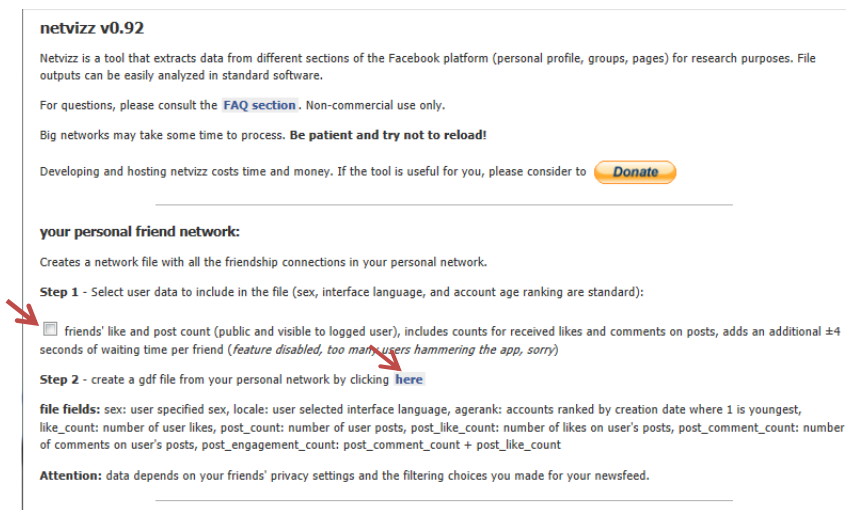
Social Network Analysis Course - Assignment 1

Facebook network

For this first assignment, you'll be working with your own Facebook network. I believe that working with your own network will make it easier to grasp concepts.

In order to get your own network, complete the following steps:

- Go to <https://apps.facebook.com/netvizz/>



netvizz v0.92

Netvizz is a tool that extracts data from different sections of the Facebook platform (personal profile, groups, pages) for research purposes. File outputs can be easily analyzed in standard software.

For questions, please consult the [FAQ section](#). Non-commercial use only.

Big networks may take some time to process. **Be patient and try not to reload!**

Developing and hosting netvizz costs time and money. If the tool is useful for you, please consider to [Donate](#)

your personal friend network:

Creates a network file with all the friendship connections in your personal network.

Step 1 - Select user data to include in the file (sex, interface language, and account age ranking are standard):

☐ friends' like and post count (public and visible to logged user), includes counts for received likes and comments on posts, adds an additional ± 4 seconds of waiting time per friend (*feature disabled, too many users hammering the app, sorry*)

Step 2 - create a gdf file from your personal network by clicking [here](#)

file fields: sex: user specified sex, locale: user selected interface language, agerank: accounts ranked by creation date where 1 is youngest, like_count: number of user likes, post_count: number of user posts, post_like_count: number of likes on user's posts, post_comment_count: number of comments on user's posts, post_engagement_count: post_comment_count + post_like_count

Attention: data depends on your friends' privacy settings and the filtering choices you made for your newsfeed.

- Select user data to include in the file, see step 1 above (sex, interface language, and account age ranking are standard). You may choose which user data (e.g. "wall posts count") you'd like to include. Note: for this assignment no additional data is necessary (meaning that you can leave the checkbox unchecked).
- Create a network file with all the friendship connections in your personal network. This is done by choosing "here" in step 2 (see red error in above image).
- Be patient. If you have many Facebook friends, the app will fetch the data gradually, and report on its progress (e.g. (30 out of 148, 120 out of 148...)). Once it is done, the page will provide a download link.
- You'll need to create a gdf file in order to load it to Gephi. Write click on the "gdf file" option to save the file locally.



- Open Gephi and load your file using "File -> Open...". Open the file as either undirected or directed network as you see fit.

If your network, like mine, has disconnected components, you will most likely want to use Force Atlas 2 as your layout algorithm. To keep the disconnected components from flying off, press "stop" or set the layout property "Gravity" to a suitable value (probably in the range of 10-50). If it is overly clumped together, you can try checking "Dissuade Hubs" and/or "Prevent Overlap". It's up to your own aesthetic sensibilities.

Hint on calculating the number of nodes in the giant component: Go to Filters (in the right section of the Gephi window). Select Topology->Giant Component. Drag the filter down where it says Queries: Drag Filter here. Then click on the 'Filter' button with the green right arrow. You should see your visualization change (only contain the giant component). Look in the upper right corner for the number of nodes and edges that are in the giant component.

Questions:

1. Did you open your network as an undirected or directed network? Why? Give an example to a network you'd like to open as the other type ("directed" if you opened it as "undirected" and vice versa).
2. Calculate the following values. These values can either be extracted from Gephi or calculated programmatically.
 - a. Number of nodes
 - b. Number of edges
 - c. Number of connected components
 - d. Size of largest connected component
 - e. Maximum degree
3. Who is the person with the highest degree in your network? Can you understand the reason (explain)
4. Node size by degree: Select Ranking tab, press the Size/Weight button, and select Degree from the drop-down menu. Set min/max sizes of Nodes, and adjust as you see fit. (If some nodes overlap each other, Run the Layout algorithm for 1 second & stop; that should fix it). Export this graph (to PDF or PNG format).
5. Community detection – apply a community detection algorithm on your network. Under Network Overview-> modularity: hit Run. Use colors to distinguish the different

- components (Partition tab on the left, select Modularity Class, and Apply). Experiment with different values of modularity resolution parameters. Export the colored network to a file in a PDF or PNG format. Examine the various communities and briefly describe the nature of each community (e.g. school friends, family, etc.)
6. Redo questions 2-4 with R code using igraph library. Note: the gml file can be read from the R code using the R function `read.graph()`. For question 4, save the plot into a PDF file (export option in the plot window), and add the file to the submission package.

Submission instructions:

Arrange all assignment results including the graphs in a file.

The assignment should be submitted to the appropriate folder on the course site, and include the following files:

1. Solutions files (document).
2. Submit graph in a GML format file (File -> Export Graph, choose GML format). The file is needed for grading.
3. R code file

The assignment due date is Sunday 23.11.2014.

Note 1: An assignment without the GML file will not be considered!!

Note 2: If privacy is a concern for you, clear the Names of your friends from the network, before you submit. This can be done if you go to Data Laboratory, and below the table, click Clear column -> Label.

Note 3: All questions regarding the assignment should be posted on the course site, in the appropriate forum. Individual mails with question on the assignment will be ignored.